Draft version February 2, 2018
Typeset using LaTeX default style in AASTeX61

POINT SPREAD FUNCTION PHOTOMETRY: THE POISSON LIKELIHOOD AND UNBIASED ESTIMATIONS

José Vinícius de Miranda Cardoso,[1,2] Geert Barentsen,[1,2] Benjamin T. Montet,[3,*] and Michael Gully-Santiago[1,2]

[1] NASA Ames Research Center, Moffett Blvd, Mountain View, CA 94035, USA
[2] Bay Area Environmental Research Institute, 625 2nd St Ste. 209, Petaluma, CA 94952
[3] Department of Astronomy and Astrophysics, University of Chicago, 5640 S. Ellis Ave, Chicago, IL 60637, USA

Keywords: point spread function — statistical estimation — photometry

Stellar photometry from charged-couple devices is at the core of astronomical observations. Often it is the case where the astronomical object of interest is isolated so that aperture photometry can be performed. In crowded regions, conversely, targets are not disjoint and therefore point spread function (PSF) photometry is the principled way to obtain meaningful measurements (Heasley 1999).

PSF photometry consists of fitting the parameters of the PSF to a given image. In many scientific analsysis, due to implications of the Central Limit Theorem (Grimmet et al. 2001), the χ^2 statistic has been traditionally and successfully used to estimate parameters.

In this letter, we advocate that the χ^2 statistic might not be adequate for PSF photometry for two main reasons: (I) the assumption of Gaussian distributed data is only compatible with the nature of pixel data in an assymptotic sense; (II) the χ^2 statistic yields a maximum likelihood estimator (MLE) for the total number of counts which is biased, and therefore it is only compatible with aperture photometry in an assymptotic sense. *ZE* TO DO: This is redundant, you could save some words here by rearranging so you don't say "in an assymptotic sense" twice

Here, we report that under Poissonian data distributed assumptions, and for a very general family of PSF models, the total number of counts obtained using PSF photometry, employing the maximum likelihood estimator, is equivalent to aperture photometry. As it will be shown, this equivalency is not valid if Gaussian distributed data is assumed.

We begin by stating the underlying statistical assumptions. Given an image as a collection of n independent *non-identically* distributed random variables $\boldsymbol{Y} \triangleq \{Y_i\}_{i=1}^{n}$ (pixels), such that $\mathbb{E}\left[Y_i\right] = \lambda_i(\boldsymbol{\Theta})$, in which λ_i is the PSF model evaluated at the i-th pixel, and $\boldsymbol{\Theta}$ is a Λ-valued random vector, where $\Lambda \subseteq \mathbb{R}^p$ is a compact set that encodes the information about, say, the total number of counts and center positions of the PSF model.

Assuming further that, for a given vector of stellar parameters $\boldsymbol{\Theta} = \boldsymbol{\theta}$, Y_i follows a Poisson distribution, then the joint conditional likelihood function of a stellar image $\boldsymbol{Y}$ can be expressed as (Grimmet et al. 2001)

$$P(\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{\Theta} = \boldsymbol{\theta}) = \prod_{i=1}^{n} p(y_i|\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{\lambda_i^{y_i}(\boldsymbol{\theta})}{y_i!} \exp -\lambda_i(\boldsymbol{\theta}) = \exp\left(-\sum_{i=1}^{n}\lambda_i(\boldsymbol{\theta})\right)\prod_{i=1}^{n}\frac{\lambda_i^{y_i}(\boldsymbol{\theta})}{y_i!}. \tag{1}$$

Perhaps of more practical interest is the log likelihood function

$$\log P(\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{\Theta} = \boldsymbol{\theta}) = \sum_{i=1}^{n}\left(-\lambda_i(\boldsymbol{\theta}) + y_i \log\lambda_i(\boldsymbol{\theta}) - \log y_i!\right). \tag{2}$$

Corresponding author: José Vinícius de Miranda Cardoso
j.demirandacardoso@nasa.gov

* Sagan Fellow

Then, the MLE can be formulated as

$$\boldsymbol{\theta}^*(\boldsymbol{y}) = \arg\min_{\boldsymbol{\theta}\in\Lambda} \sum_{i=1}^{n} \left(\lambda_i(\boldsymbol{\theta}) - y_i \log \lambda_i(\boldsymbol{\theta})\right), \tag{3}$$

which must satisfy the following system of differential equations

$$\sum_{i=1}^{n} \frac{\partial \lambda_i(\boldsymbol{\theta})}{\partial \theta_j} \left(1 - \frac{y_i}{\lambda_i(\boldsymbol{\theta})}\right)\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*(\boldsymbol{y})} = 0, \text{ for } j = 1, 2, ..., p, \tag{4}$$

assuming that $\boldsymbol{\theta}^*(\boldsymbol{y})$ is an interior point of Λ.

Now, assume that the vector of stellar parameters $\boldsymbol{\theta}$ takes the following form $\boldsymbol{\theta} \triangleq (\alpha, \boldsymbol{x})$, where α is the total number of counts and $\boldsymbol{x} \triangleq (x_1, x_2)$ is the coordinate of the center of the PSF. Further, assume that the PSF model can be factorized as follows

$$\lambda_i(\boldsymbol{\theta}) = \alpha \cdot \psi_i(\boldsymbol{x}), \tag{5}$$

where $\psi_i(\boldsymbol{x})$ is the normalized PSF model centered at $\boldsymbol{x}$ at the i-th pixel, such that $\sum_{i=1}^{n} \psi_i(\boldsymbol{x}) = 1$.

Therefore,

$$\frac{\partial \lambda_i(\boldsymbol{\theta})}{\partial \alpha} = \psi_i(\boldsymbol{x}) \text{ and } \frac{\partial^2 \lambda_i(\boldsymbol{\theta})}{\partial \alpha^2} = 0. \tag{6}$$

Substituting (5) and (6) into (4), it follows that

$$\sum_{i=1}^{n} \psi_i(\boldsymbol{x}^*) \left(1 - \frac{y_i}{\alpha^* \cdot \psi_i(\boldsymbol{x}^*)}\right) = \sum_{i=1}^{n} \left(\psi_i(\boldsymbol{x}^*) - \frac{y_i}{\alpha^*}\right) = 0 \quad \Leftrightarrow \quad \alpha^* = \sum_{i=1}^{n} y_i. \tag{7}$$

Furthermore, computing the derivative of the left hand side of (4), *i.e.*, the second-order derivative of the negative of the log likelihood function, with respect to α, and evaluating it at α^*, it follows that

$$\sum_{i=1}^{n} \left\{ \frac{\partial^2 \lambda_i(\boldsymbol{\theta})}{\partial \alpha^2} - \frac{y_i}{\lambda_i(\boldsymbol{\theta})} \left[\frac{\partial^2 \lambda_i(\boldsymbol{\theta})}{\partial \alpha^2} - \frac{1}{\lambda_i(\boldsymbol{\theta})} \left(\frac{\partial \lambda_i(\boldsymbol{\theta})}{\partial \alpha} \right)^2 \right] \right\} \Big|_{\alpha=\alpha^*} = \sum_{i=1}^{n} \frac{y_i}{\alpha^2}\Big|_{\alpha=\alpha^*} = \frac{1}{\alpha^*} > 0 \tag{8}$$

which proves our proposition. Note that this result does not require any specific structural form of λ_i other than it increases linearly with the total number of counts α for a given fixed position $\boldsymbol{x}$.

Furthermore, the MLE for the total number of counts has the following nice properties: (I) it is an unbiased estimator and (II) its variance attains the minimum possible bound (Cramér-Rao Lower Bound (Bobrovsky et al. 1987)) in the Fisher information sense. The first property may be verified by taking the expectation value of (7), which follows that

$$\mathbb{E}[\alpha^*] = \sum_{i=1}^{n} \mathbb{E}[Y_i] = \sum_{i=1}^{n} \lambda_i(\boldsymbol{\theta}) = \sum_{i=1}^{n} \alpha \cdot \psi_i(\boldsymbol{x}) = \alpha. \tag{9}$$

To verify the second property, note that α^* is the sum of independent Poisson random variables, therefore, $var[\alpha^*] = \mathbb{E}[\alpha^*] = \alpha$.

Note that, under the widely adopted assumption that the data is Gaussian distributed (corresponding to a χ^2 likelihood function), the solution for α^* would be

$$\alpha^* = \frac{\sum_{i=1}^{n} y_i \psi_i(\boldsymbol{x}^*)}{\sum_{i=1}^{n} \psi_i^2(\boldsymbol{x}^*)}, \tag{10}$$

for which, properties (I) and (II) are not known to hold.

It is worth noticing that, whenever the number of counts is high (which often is the case around the center of the PSF), the Poisson density does converge to a Gaussian density, and therefore those two approaches are equivalent.

However, for pixels at the tail of the PSF or low signal-to-noise targets, where the number of counts is small, the Gaussian approximation might yield residuals that are nonnegligeble for practical scenarios such as transit photometry.

As a practical exercise, we apply PSF photometry on the target `KIC10053146` of NASA's Kepler Mission. `KIC10053146` is a 20th magnitude isolated star which has an average flux of $260e^{-}s^{-1}$ in the brightest pixel.

Using the assumption that the data is Poisson distributed, we obtain a mean residual value at the two brightest pixels across 4000 cadences of $9.35 \pm 4.45\ e^{-}s^{-1}$. The same analysis with Gaussian likelihood assumption yields $10.11 \pm 5.76\ e^{-}s^{-1}$. The mean residuals considering all 30 pixels was $-2.25 \cdot 10^{-5} \pm 12.58\ e^{-}s^{-1}$, and $0.99 \pm 12.59\ e^{-}s^{-1}$ for the Poisson and Gaussian assumptions, respectively.

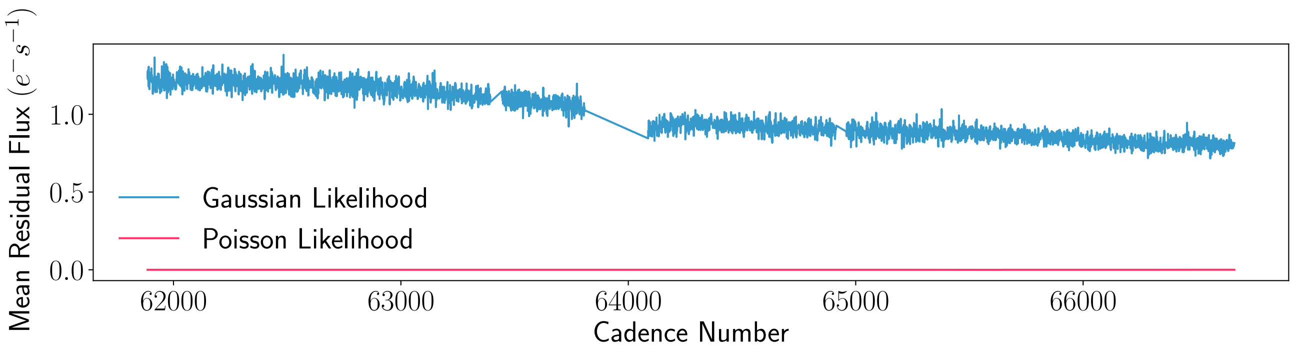

Figure 1. Mean residuals per cadence of PSF photometry using Poisson Likelihood (red) and Gaussian Likelihood (blue).

This work acknowledges the use of the following computational tools: `numpy` (van der Walt et al. 2011), `scipy` (Jones et al. 2001), `astropy` (AstroPy Project 2017), `pyke` (PyKE Contributors 2017), and `jupyter` (Perez et al. 2007). A `jupyter` notebook intented to reproduce the example aforementioned is available at: https://www.github.com/mirca/geerts-conjecture. Work by B.T.M. was performed under contract with the Jet Propulsion Laboratory (JPL) funded by NASA through the Sagan Fellowship Program executed by the NASA Exoplanet Science Institute.

REFERENCES

Heasley, J. N., Point-Spread Function Fitting Photometry, 1999, Astronomical Society of the Pacific Conference Series.

Grimmett, R. G., Stirzaker D. R., Probability and Random Processes, 2001, Oxford University Press.

Bobrovsky, B. Z., Mayer-Wolf E., Zakai M., Some Classes of Global Cramér-Rao Bounds, 1987, The Annals of Statistics, Vol. 15, No. 4, 1421 – 1438.

Jones, E., Oliphant T., Peterson P., et al., `SciPy`: Open source scientific tools for `Python`.

van der Walt S., Colbert S. C., Varoquaux G., The NumPy Array: a structure for efficient numerical computation, 2011, Computing in Science & Engineering, 13, 22–30.

Hunter, J. D., Matplotlib: A 2D graphic environment, 2007, Computing in Science & Engineering, 9, 90–95.

AstroPy Project, Submitted to ApJ, 2017.

PyKE Contributors, PyKE: Kepler, *K2* & TESS Data Analysis Tools, 2017, http://doi.org/10.5281/zenodo.835584.

Pérez F., Granger B. E., `IPython`: a system for interactive scientific computing, 2007, Computing in Science and Engineering, 9, 21–29.